Engineering the S1′ Subsite of Trypsin: Design of a Protease Which Cleaves between Dibasic Residues†

Torsten Kurth,‡,§ Sibylla Grahn,‡ Michael Thormann,‡ Dirk Ullmann,‡,|| Hans-Jörg Hofmann,‡ Hans-Dieter Jakubke,*,‡ and Lizbeth Hedstrom*,§

Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02454, and Institute of Biochemistry, Faculty of Biosciences, Pharmacy, and Psychology, University of Leipzig, 04103 Leipzig, Germany

Received April 15, 1998

ABSTRACT: The serine protease trypsin was converted into a site-specific protease which hydrolyzes peptides between dibasic residues. Trypsin exhibits a high S1 specificity for Arg and Lys residues. However, the S1′ specificity of trypsin is very broad, with only a slight preference for hydrophobic residues in P1′. We replaced Lys60 with Glu and Asp to introduce a high specificity for basic residues into the S1′ site of trypsin. Both mutations cause a dramatic increase in the S1′ specificity for Arg and Lys as measured by acyl transfer reactions. In K60E, the preference for Arg increases 70-fold while the preference for P1′-Lys increases 12-fold. In contrast, the preferences for other P1′ residues either decrease slightly or remain the same. Thus, K60E prefers P1′-Arg over most other P1′ residues by 2 orders of magnitude. Similar results are obtained when P1′ specificity is measured in peptide cleavage assays. K60D exhibits an S1′ specificity profile very similar to that of K60E, although the P1′-Arg preference is reduced by a factor of 2.5. Molecular modeling studies suggest that the high S1′ specificity for Arg in K60E may be due to the formation of a salt bridge between Glu60 and the P1′-Arg of the substrate.

The design of site-specific proteases is a challenging protein engineering problem of both scientific and economic importance. This field can claim few successes in the remodeling of serine proteases: (i) the redesign of subtilisin by comparison with kex2 and furin to yield a protease which recognizes basic residues in S1, S2, and S4 (*1*−*3*; nomenclature according to Schechter and Berger), (ii) the introduction of specficity for His into S2 and S1′ of subtilisin and trypsin by "substrate-assisted" catalysis (*4*, *5*); and (iii) the use of a metal binding site to select for His residues at S2′ in trypsin (*6*). We now report the redesign of the S1′ site of trypsin to create a protease which cleaves between dibasic residues.

Although trypsin specificity is primarily determined by the preference of the S1 site for Arg and Lys residues, the adjacent binding sites are also important for efficient substrate hydrolysis. For example, trypsin hydrolyzes oligopeptide substrates 10^3-fold more efficiently than single-amino acid substrates. Trypsin does not possess carboxypeptidase activity, which further illustrates the importance of the adjacent subsites (*7*). Similarly, trypsin catalyzes acyl transfer to peptide nucleophiles and amino acid amides but not to free amino acids (*8*). Nevertheless, the S′ specificity of trypsin is very broad. Trypsin prefers substrates containing aliphatic residues at P1′, but will also utilize aromatic and, to a lesser extent, positively charged P1′ residues (*9*, *10*). Only P1′-Pro residues are not tolerated in S1′ of trypsin. The S2′ site of trypsin also has broad specificity, with a slight preference for Arg and Lys residues (*10*). The S3′ site overlaps the S1′ site and, therefore, has a specificity similar to that of the S1′ site.

We have initiated a program for investigating how structure determines function in the trypsin family of serine proteases. Previously, we changed trypsin into a protease which prefers substrates containing large hydrophobic residues at P1. The exchange of two surface loops with the analogous loops of chymotrypsin is critical for this specificity change (*11*, *12*). Recently, we extended this work to the S1′ site, demonstrating that the exchange of two different surface loops (loop 40, residues 33−41, and loop 60, residues 58−68) with the analogous residues of chymotrypsin transfers chymotrypsin's specificity for P1′-Arg and/or Lys (*13*). The relative P1′ specificity for Arg versus Met changes from 0.2 in trypsin to 2.2 in the mutant Tr → Ch′(L40+L60), as compared to 4−7 in chymotrypsin. Although this is a significant change in the relative specificities of P1′-Met and Arg, the overall substrate discrimination is not impressive.

In this work, we utilize a different strategy to introduce specificity for P1′-Arg into trypsin. Molecular modeling studies revealed that the side chain of Lys60 is the only side chain of loop 40 and loop 60 directly exposed to the P1′ residue. The introduction of Glu at position 60 in trypsin creates a protease that specifically cleaves between adjacent Arg residues.

† This work was supported by the German Scholarship Foundation (Studienstiftung des deutschen Volkes) (T.K.), an NSF Career Award (L.H.), and the Deutsche Forschungsgemeinschaft (INK 23/A2). L.H. is a Searle Scholar and a Beckman Young Investigator. This work is publication no. 1821 from the Department of Biochemistry, Brandeis University.

* To whom correspondence should be addressed.

‡ University of Leipzig.

§ Brandeis University.

|| Present address: EVOTEC Biosystems, Grandweg 64, Hamburg 22529, Germany.

S0006-2960(98)00842-3 CCC: $15.00
Published on Web 07/31/1998

MATERIALS AND METHODS

Construction of Trypsin Mutants. Site-directed mutagenesis was performed by the method of Kunkel as described previously (*12*, *14*). All mutants were completely sequenced to ensure that only the desired mutation was introduced. The following oligonucleotides were used (mismatches are underlined): (1) K60D, GCT CAC TGC TAT GAT TCC CGC ATC CAA; and (2) K60E, GCT CAC TGC TAT GAA TCC CGC ATC CAA. Trypsin and trypsin mutants were isolated and purified as described previously (*12*).

Acyl Transfer Experiments. Reactions were performed at 25 °C. A 4 mM stock solution of the acyl donor Bz-Arg-OEt[1] in water was prepared daily. Stock solutions of the amino acid amides and pentapeptides (each 30 mM) were prepared in assay mix and readjusted with NaOH to pH 8. The total assay volume was 65 μL. The final acyl donor concentration was 2 mM, and the nucleophile concentrations were 15 mM which were calculated as unprotonated amino acid amide concentration [N] according to the following equation ($[N]^O$ corresponds to the total nucleophile concentration):

$$[N] = [N]^O/(1 + 10^{pK-pH}) \quad (1)$$

The acyl transfer reaction was initiated with 5 μL of enzyme stock solution. The enzyme concentration and reaction time were adjusted to achieve an ester consumption of 50−80% to ensure that no secondary hydrolysis of the peptide product occurs. The reaction was stopped by diluting 50 μL of the reaction mixture to 0.3 mL of 50% aqueous methanol and 1% trifluoroacetic acid. The partition values were determined from three to five independent experiments. Control experiments without the enzyme were performed to estimate nonenzymatic hydrolysis (0−1.5%). HPLC was performed using a Thermoseparation product system on a Vydac analytical reversed phase C_{18} column (Vydac TP218 54). Samples were eluted under isocratic conditions with eluents containing 15−25% acetonitrile (depending on the nucleophile) in 0.1% aqueous trifluoroacetic acid at flow rates of 1.0−1.2 mL/min. Absorbance was monitored at 254 nm. The ratio between aminolysis and hydrolysis product was calculated from the corresponding peak areas. The ratio of absorption coefficients of the sample components was obtained using hydrolysis experiments according to Ullmann and Jakubke (*15*).

Synthesis of Peptide Substrates. The peptide substrates were synthesized using standard Fmoc solid phase protocols with a semiautomatic batch peptide synthesizer SP650 (Labortec) and *p*-alkoxybenzyl alcohol resin. The peptides were deprotected and removed from the resin with TFA/scavenger cocktail, precipitated with diethyl ether, and purified on a Sephadex LH-20 column. The purity of all peptides was determined by analytical HPLC [Vydac 218TP54 column, elution with an acetonitrile/water gradient and UV detection (220 nm)]. Molecular mass determination was performed with MALDI-TOF analysis: GPRRPAAG, 781.45 (M + H); GPRMPAAG, 756.40 (M + H); GPRFPAAG, 772.36 (M + H); and GPRAPAAG, 696.43 (M + H).

Synthesis of FRET Substrates. Attachment of the DABCYL group to the peptide resin and the introduction of the EDANS group were performed as previously described (*16*). The protecting groups (Lys-Z) and (Arg-Mts) were removed in a final step after the coupling of the fluorescence donor and the removal of the product from the resin. The purity and molecular mass were determined as described above (MALDI-TOF-MS): DABCYL-GPAKLAAG-EDANS, 1248.90 (M + H); DABCYL-GPARLAAG-EDANS, 1276.48 (M + H); and DABCYL-GPRRPAAG-EDANS, 1281.64 (M + H).

FRET Assays. Fluorometric assays were performed in a 1 mL reaction volume containing 50 mM Hepes (pH 8.0), 10 mM $CaCl_2$, 0.1 M NaCl, and 0.5−25 μM DABCYL-EDANS-labeled substrates for enzyme hydrolysis studies. Stock solutions of the internally quenched substrates were prepared in methanol and stored at 4 °C. The final concentration of methanol in the assay was less than 5%. The competition experiments contained 20 μM DABCYL-Gly-Pro-Arg-Arg-Pro-Ala-Ala-Gly-EDANS, and 0.02−10 mM competetive substrates Gly-Pro-Arg-Xaa-Pro-Ala-Ala-Gly were used. Each reaction was initiated by addition of enzyme. The fluorescence was measured with a Perkin-Elmer LS50B luminescence spectrometer at an excitation wavelength of 336 nm and an emission wavelength of 490 nm. Concentrations of fluorogenic substrates were determined by using a standard curve of fluorescence of the complete enzyme hydrolysate of the substrates at 12 different concentrations.

The steady-state kinetic parameters k_{cat}, K_M, and k_{cat}/K_M were calculated employing the nonlinear regression data analysis software EnzFitter (Elsevier Biosoft). Competitive experiments were analyzed with linear regression of v_0/v against [C]/[S] according to the following equation:

$$v_0/v = 1 + (k_{cat}/K_M)_C[C]/(k_{cat}/K_M)_S[S] \quad (2)$$

where C denotes the competitor peptide, S denotes the substrate, and v_0 denotes the rate of substrate hydrolysis at [C] = 0.

Molecular Modeling Studies. The modeling investigations are based on the X-ray structure of rat trypsin (D189S/G226D) complexed with BPTI (PDB code 1brb; *17*). The Quanta96 modeling program package (Quanta96, Molecular Simulations, Inc., San Diego, CA) was used to build the complex of the peptide substrate Bz-Arg-Arg-NH_2 with trypsin by fitting the arginines onto Lys15-Ala16 of BPTI. Lys60 was then replaced by Glu to obtain the mutant complex. After addition of the polar hydrogens, template charges were assigned to the two proteins, resulting in total charges of −7 for the wild type and −9 for K60E. To allow some local conformational adaptation for Glu60, a constrained gas phase dynamics was performed keeping the protein fixed except for residues 58−62. During an evolution time of 100 ps, 500 samples were registered and reoptimized. Glu60 forms a salt bridge to the P1′-Arg of the peptide in the most stable conformation found by this method. This protein conformation was the basis for the subsequent investigations applying a Monte Carlo simulated annealing docking procedure using the program AutoDock 2.4 (*18*). To search for the best conformations for binding of the

[1] Abbreviations: DABCYL, 4-[[4-(dimethylamino)phenyl]azo]benzoic acid; EDANS, 5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid; FRET, fluorescence resonance energy transfer; Fmoc, fluorenyl-9-methyloxycarbonyl; Mts, mesystilensulfonyl; Z, benzyloxycarbonyl.

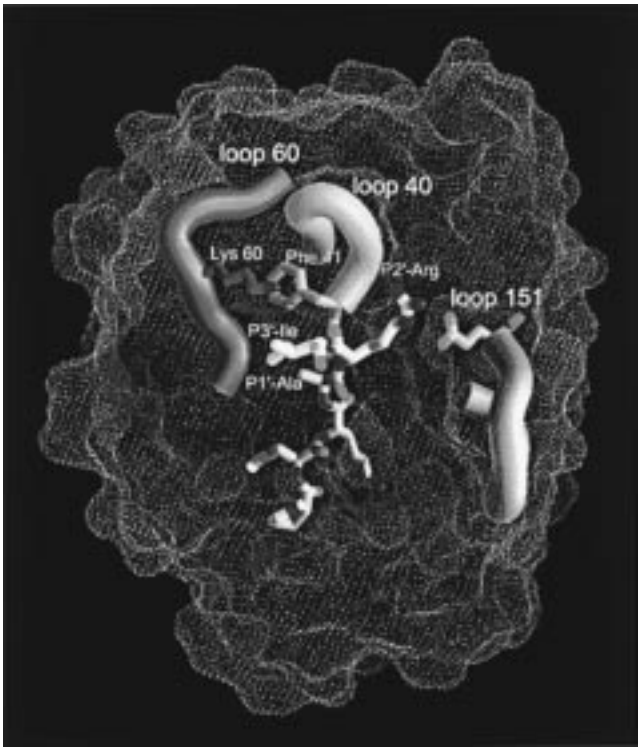

FIGURE 1: S′ subsites of trypsin. The structure of the rat trypsin (G226D/D189S)−BPTI complex (Brookhaven Protein Data Bank identification number 1brb) has been processed with Grasp. Three loops (blue, loop 60 from residue 57 to 64; orange, loop 40 from residue 33 to 41; and green, loop 151 from residue 143 to 151) and residues 13−19 of the BPTI inhibitor are shown in their position relative to the S′ subsite loops. The protein surface is dotted with negatively charged regions in red and positively charged in blue.

peptide Bz-Arg-Arg-Ala-NH_2 to the wild type and K60E, a cubic grid 25 Å in length was centered at Ser195 with a grid spacing of 0.25 Å according to a described formalism (*19*). In the starting position of the docking procedure, neither the P1-Arg nor the P1′-Arg fits into the S1 subsite or contacted Glu60. All torsion angles, except those for the peptide and the guanidinyl group bonds, were unconstrained in the docking runs. Sixty million protein−ligand conformations were computed which resulted from 20 runs with 120 cycles per run and 25 000 steps per cycle applying the same temperature schedule. A distance-dependent permittivity was applied in all calculations. The Grasp 1.3.6 (*20*) and the Ligplot 3.2 (*21*) programs were used to visualize the most stable complexes calculated. Calculations for the binding conformations of Bz-Arg-Lys-Ala-NH_2 to the wild type and K60E were performed analogously.

RESULTS AND DISCUSSION

Mutant Design. We evaluated the available structure information in order to design trypsin mutants with high substrate specificity in the S′ subsites. The X-ray crystal structure of the rat trypsin (D189S/G226D)−BPTI complex suggests that few enzyme−substrate interactions are formed in the S′ subsites, which might explain the broad specificity in this region (*17*). Additionally, molecular modeling studies and site-directed mutagenesis experiments suggest that residues of three surface loops comprise the S1′−S3′ subsites of trypsin: loop 40, loop 60, and loop 151 (Figure 1; *13*). Three notable enzyme−inhibitor interactions are observed in this region. (1) A hydrogen bond is formed between the carbonyl of Phe41 and the amide group of the P2′ residue. The absence of this hydrogen bond may account for the absence of carboxypeptidase activity in trypsin. (2) A salt bridge is formed between Glu151 and P2′-Arg. This salt bridge can account for the S2′ preference for positively charged residues. (3) Lys60 is most exposed to the P1′ and P3′ residues of the inhibitor. Therefore, Lys60 could be an important determinant of S1′ and/or S3′ specificity in trypsin. Consistent with this hypothesis, the S1′ site has a low specificity for positively charged residues, which can be attributed to unfavorable charge interactions between P1′-Lys and Arg residues and Lys60 (*10*). In addition, the preference for hydophobic residues in P1′ and/or P3′ might result from interactions with the alkyl side chain of Lys60. Therefore, substitutions at Lys60 seem to be a promising strategy for creating a new S1′ specificity. We constructed two trypsin mutants, K60E and K60D. Both enzymes were expressed in a *Saccharomyces cerevisae* system and purified using standard methods (*12*).

Table 1: Steady-State Kinetic Parameters for the Hydrolysis of FRET Substrates by Trypsin and Trypsin Mutants[a]

enzyme	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
DABCYL-Gly-Pro-Arg↓Leu-Ala-Ile-Gly-EDANS			
trypsin	34 ± 3	40 ± 3	(1.2 ± 0.2) × 10^6
K60D	12 ± 1	2 ± 0.1	(1.7 ± 0.2) × 10^5
K60E	53 ± 6	9 ± 0.8	(1.7 ± 0.3) × 10^5
DABCYL-Gly-Pro-Lys↓Leu-Ala-Ile-Gly-EDANS			
trypsin	58 ± 5	10 ± 0.6	(1.8 ± 0.2) × 10^5
K60D	nd	nd	(2.3 ± 0.04) × 10^4
K60E	nd	nd	(1.8 ± 0.01) × 10^4
DABCYL-Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly-EDANS			
trypsin	3.7 ± 0.2	0.8 ± 0.02	(2.3 ± 0.2) × 10^5
K60D	3.6 ± 0.3	5 ± 0.2	(1.4 ± 0.2) × 10^6
K60E	2.8 ± 0.2	8 ± 0.2	(2.8 ± 0.2) × 10^6

[a] Conditions: 50 mM Hepes (pH 8), 10 mM $CaCl_2$, and 0.1 M NaCl at 25 °C with a λ_{ex} of 336 nm and a λ_{em} of 490 nm.

Kinetic Characterization of K60E and K60D. We utilized FRET substrates to evaluate the basic kinetic parameters of peptide cleavage catalyzed by K60D and K60E. While fluorescence is quenched in the intact substrate, cleavage of the substrate leads to an increase of fluorescence intensity (*22*, *23*). Thus, FRET substrates allow the continuous monitoring of peptide cleavage. Three different substrates were synthesized. DABCYL-Gly-Pro-Arg↓Leu-Ala-Ile-Gly-EDANS and DABCYL-Gly-Pro-Lys↓Leu-Ala-Ile-Gly-EDANS were based on known trypsin substrates. In addition, DABCYL-Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly-EDANS was synthesized to include P1′-Arg in the peptide sequence. Here we introduced Pro into P2′ to ensure that no cleavage occurs at the second Arg residue.

As shown in Table 1, trypsin prefers P1-Arg over P1-Lys by 1 order of magnitude, as observed previously (*24*). K60D and K60E also prefer P1-Arg over P1-Lys. Thus, these mutations did not substantially alter the S1 specificity of trypsin. However, the values of k_{cat}/K_M for the hydrolysis of DABCYL-Gly-Pro-Arg↓Leu-Ala-Ile-Gly-EDANS and DABCYL-Gly-Pro-Lys↓Leu-Ala-Ile-Gly-EDANS catalyzed by K60D and K60E are 10-fold lower than the equivalent values for trypsin. In contrast, both mutants catalyze the cleavage of DABCYL-Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly-EDANS more efficiently than trypsin (by 1 order of

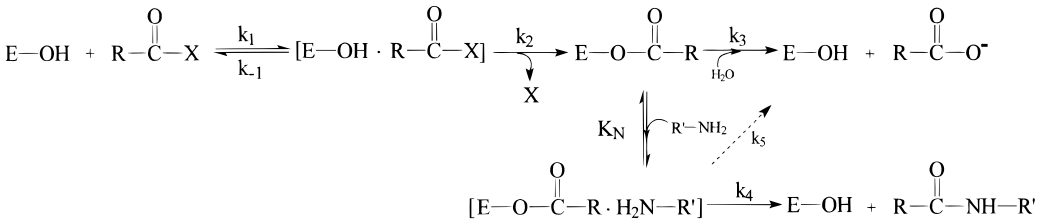

FIGURE 2: Mechanism of protease-catalyzed acyl transfer reactions: E-OH, enzyme; R-COX, acyl donor; X, leaving group; and R′-NH_2, nucleophile.

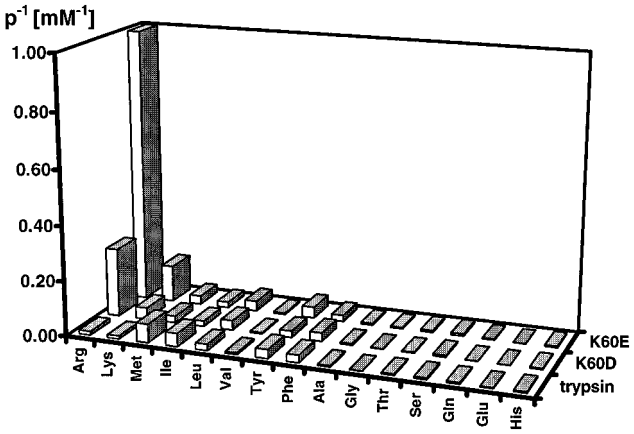

FIGURE 3: S_1' specificity of trypsin and the mutant trypsins K60D and K60E as determined with acyl transfer reactions from Bz-Arg-OEt on amino acid amide nucleophiles. Assays were performed at 25 °C in assay mix [50 mM Hepes (pH 8), 10 mM $CaCl_2$, 100 mM NaCl, and 0.01% Triton] containing 2 mM BAEE and 15 mM R-NH_2. Reactions were started with enzyme. Reactions were stopped by adding 50% methanol and 1% TFA after 50−80% of the ester had been consumed. Products were analyzed via reversed phase HPLC (Vydac CAT 218TP54, 12−20% acetonitrile, $\lambda = 254$ nm). The partition value was calculated with the equation $p = [\text{Ac-OH}][\text{R-NH}_2]/[\text{Ac-NH-R}]$. Since [R-$NH_2$] ≫ [BAEE], it can be assumed that [R-NH_2] remains constant during the course of the reaction.

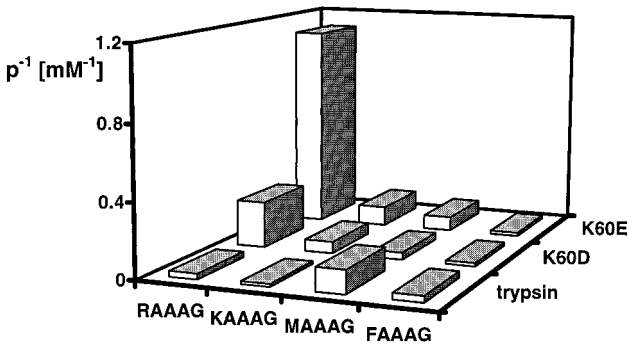

FIGURE 4: S_1' specificity of trypsin, trypsin K60D, and trypsin K60E as determined with acyl transfer reactions from Bz-Arg-OEt on pentapeptide nucleophiles. Assays were perfomed as described in Figure 3.

magnitude in k_{cat}/K_M). These observations suggest that the mutation of Lys60 changes the S1′ specificity.

Determination of S1′ Specificity. The S′ specificity of serine and cysteine proteases can be determined by monitoring acyl transfer reactions with added nucleophiles (*25*). This reaction is the reverse of peptide hydrolysis and, therefore, provides analogous specificity data (Figure 2; *8*, *9*). Since the acyl group of a substrate R-COX, which is usually an ester, can be transferred to both water and R′-NH_2, two products are formed: R-COOH and R-CO-NHR′, respectively. The ratio between hydrolysis and aminolysis is determined by the S′ specificity of the protease. The partition value *p* describes the acyl transfer efficiency and therefore the S′ specificity of a protease. *p* is defined according to Figure 2:

$$p = [\text{R'-NH}_2]v_H/v_A = k_3K_N/k_4 \quad (3)$$

where v_H and v_A represent the velocities of hydrolysis and aminolysis, respectively. The partition value can be determined from the product ratios when R′-NH_2 is in excess:

$$p = [\text{R'-NH}_2]_0[\text{R-CO}_2\text{H}]/[\text{R-CONH-R'}] \quad (4)$$

where $[\text{R'-NH}_2]_0$ is the initial nucleophile concentration and [R-CO_2H] and [R-CONH-R′] represent the product concentrations. Consequently, the value 1/*p* can be correlated

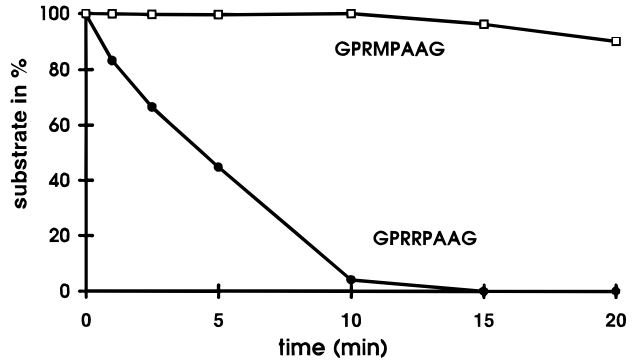

FIGURE 5: Cleavage of Gly-Pro-Arg-Arg-Pro-Ala-Ala-Gly in the presence of a 10-fold molar excess of Gly-Pro-Arg-Met-Pro-Ala-Ala-Gly by the trypsin variant K60E. Conditions were as follows: 50 mM Hepes (pH 8), 10 mM $CaCl_2$, and 0.1 M NaCl at 25 °C with UV irradiation (220 nm). [GPRRPAAG] = 20 μM, and [GPRMPAAG] = 200 μM.

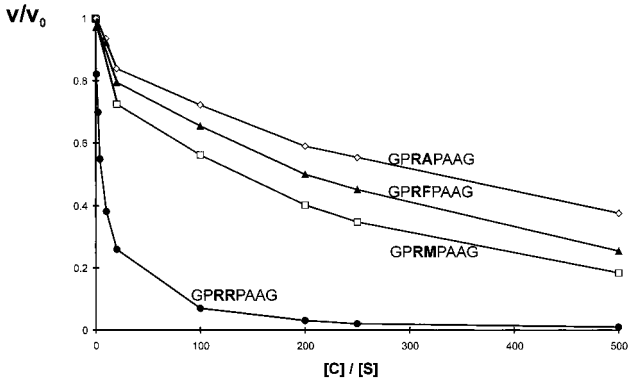

FIGURE 6: Inhibition of hydrolysis of DABCYL-Gly-Pro-Arg-Arg-Pro-Ala-Ala-Gly-EDANS by substrates competing for the trypsin variant K60E. Conditions were as follows: 50 mM Hepes (pH 8), 10 mM $CaCl_2$, and 0.1 M NaCl at 25 °C with a λ_{ex} of 336 nm and a λ_{em} of 490 nm. [DABCYL-GPRRPAAG-EDANS] = 20 μM, and [GPR-Xaa-PAAG] = 0.02−10 mM.

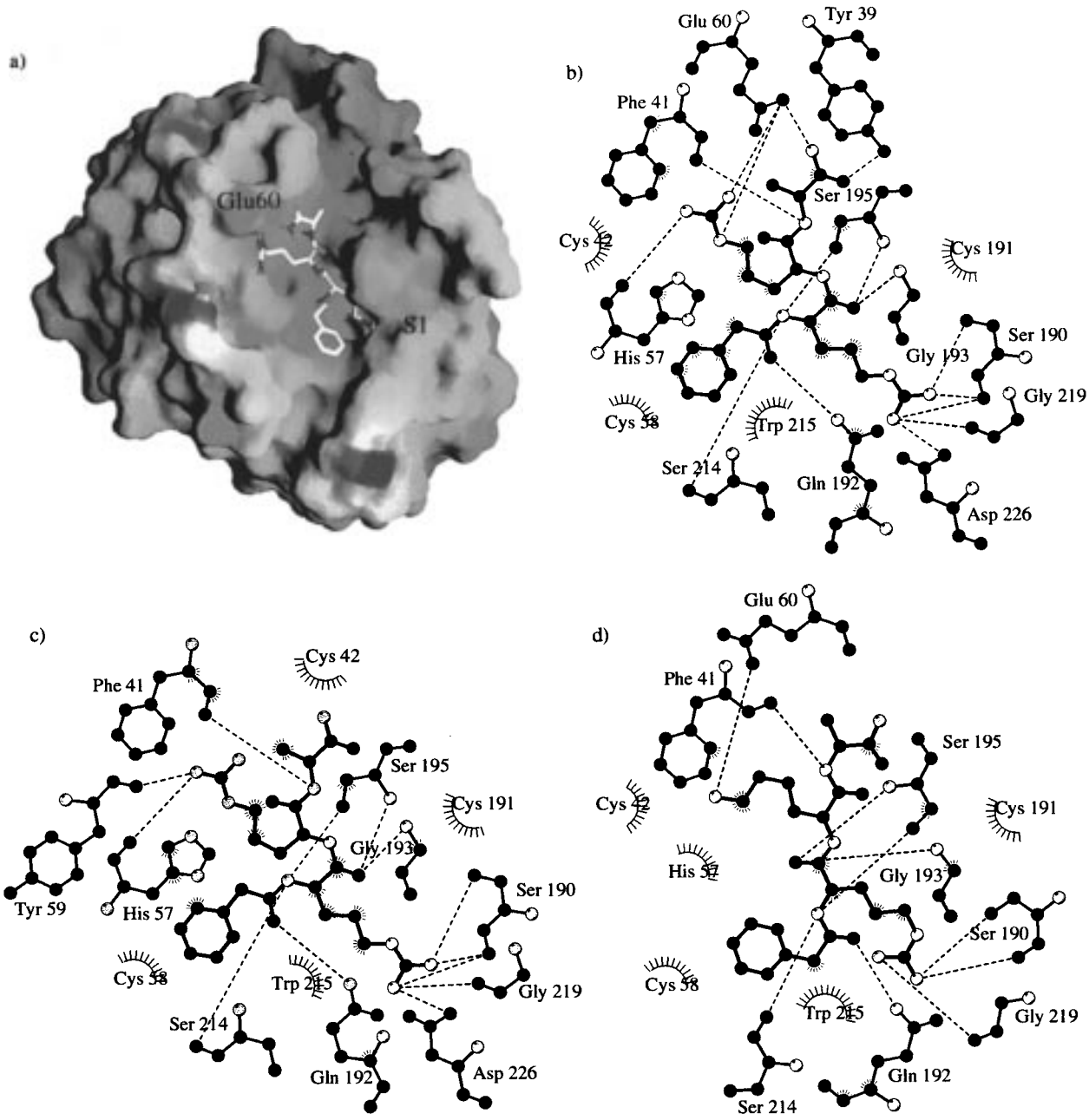

FIGURE 7: Structural basis of S1′ specificity of trypsin and trypsin K60E as calculated via molecular modeling. The surface of the trypsin K60E mutant is shown with the lowest-energy conformers of the Bz-Arg-Arg-Ala-NH_2 peptide resulting from the AutoDock procedure (a). Positively charged surface regions are blue, while negatively charged regions are red. To facilitate the identification of the peptide−protein interactions, the same complex is depicted as a Ligplot drawing (b). Additionally, Ligplot drawings of the complexes wild-type-Bz-Arg-Arg-Ala-NH_2 (c) and K60E-Bz-Arg-Lys-Ala-NH_2 (d) are depicted. The dotted lines show hydrogen bonds between the ligand and the protein residues. The van der Waals contacts are also shown.

directly with the preference of the protease for a given R′-NH_2. Therefore, the S1′ specificities of trypsin and the mutant enzymes were determined by measuring $1/p$ for 15 amino acid amides and four pentapeptide nucleophiles using Bz-Arg-OEt as the acyl donor.

S1′ Specificity of Trypsin, K60E, and K60D. The mutation of Lys60 dramatically changes the S1′ specificity of trypsin (Figures 3 and 4). The $1/p$ value of K60E for the acyl transfer on H-Arg-NH_2 and the pentapeptide H-Arg-$(Ala)_3$-Gly is increased by nearly 2 orders of magnitude compared to that of trypsin. The analogous values for P1′-Lys are increased by factors of 12 and 4. The $1/p$ values for the other nucleophiles either decrease slightly or remain the same so that K60E has a very high selectivity for P1′-Arg. This is also manifest in the P1′-Arg versus Met discrimination, which has a value of 30 for K60E compared to 0.20 for trypsin and 2.2 for the previously described mutant Tr′ → Ch′(L40+L60). Thus, the trypsin mutant K60E prefers P1′-Arg over all other residues, with the exception of P1′-Lys, by at least 1 order of magnitude and over most nucleophiles by 2 orders of magnitude. The results from acyl transfer reactions on amino acid amides can be confirmed with pentapeptides of the structure H-Xaa-$(Ala)_3$-Gly-OH (Figure 4).

K60D exhibits an S1′ specificity profile similar to that of K60E, but the S1′ specificity for Arg is decreased by a factor of 2.5. Nevertheless, K60D still discriminates between P1′-Arg and P1′-Met by a factor of 12 and, therefore, exhibits a

relatively strong S1′ specificity for Arg. Thus, both mutations of Lys60 create a very specific Arg binding site in S1′. Further, these results demonstrate that residue 60 is the major determinant of S1′ substrate specificity.

K60E Is Specific for the Cleavage of Arg−Arg Bonds. To confirm the results from acyl transfer experiments, we measured the specificity of K60E in peptide hydrolysis assays. In the first experiment, a mixture of Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly and a 10-fold molar excess of Gly-Pro-Arg↓Met-Pro-Ala-Ala-Gly was incubated with K60E. The cleavage of both substrates was monitored by HPLC. As shown in Figure 5, the P1′-Arg-containing peptide was hydrolyzed by K60E exclusively and no turnover of Gly-Pro-Arg↓Met-Pro-Ala-Ala-Gly was observed. Both substrates were cleaved with comparable rates in an analogous experiment with trypsin.

To further demonstrate the selectivity of K60E for Arg−Arg bonds, we monitored the hydrolysis of DABCYL-Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly-EDANS in the presence of four competing nonfluorogenic peptides of the general structure Gly-Pro-Arg-Xaa-Pro-Ala-Ala-Gly (Xaa is Arg, Met, Phe, or Ala). A saturating concentration of DABCYL-Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly-EDANS was utilized so that hydrolysis of the competitors results in an equivalent decrease in the rate of hydrolysis of the FRET substrate.

As shown in Figure 6, a 10-fold molar excess of the P1′-Met-, Phe-, and Ala-containing peptides does not significantly decrease the rate of hydrolysis of DABCYL-Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly-EDANS, consistent with the first experiment. Inhibition is observed at higher concentrations of these competitors, and when the competitor contains P1′-Arg. The ratio of $(k_{cat}/K_M)_S/(k_{cat}/K_M)_C$ can be determined from eq 2. This ratio is 4.5 for the P1′-Arg competitor. This ratio should equal 1 if the competitor and DABCYL-Gly-Pro-Arg↓Arg-Pro-Ala-Ala-Gly-EDANS are equivalent substrates. This observation suggests that the DABCYL and/or EDANS moieties interact with trypsin. Nevertheless, the ratios of $(k_{cat}/K_M)_S/(k_{cat}/K_M)_C$ for the P1′-Met, P1′-Phe, and P1′-Ala competitors correlate very well with the ratios of p determined by acyl transfer experiments: P1′-Met, $(k_{cat}/K_M)_S/(k_{cat}/K_M)_C = 132$ versus 32 using p; P1′-Phe, $(k_{cat}/K_M)_S/(k_{cat}/K_M)_C = 170$ versus 44; and P1′-Ala, $(k_{cat}/K_M)_S/(k_{cat}/K_M)_C = 400$ versus 170. Thus, the high P1′-Arg specificity of K60E is also evident in peptide cleavage experiments.

Structural Basis of S1′ Specificity in K60E. Since the preference of K60E is much stronger for P1′-Arg than for P1′-Lys, we believed that the P1′-Arg might specifically interact with Glu60. To further investigate this hypothesis, we modeled the interaction between K60E and Bz-Arg-Arg-Ala-NH_2 on the basis of the structure of the rat trypsin (D189S/G226D)−BPTI complex. To determine the best binding conformations, we performed docking calculations with trypsin, K60E, and Bz-Arg-Arg-Ala-NH_2. The most stable structures found for the two protein complexes are depicted in Figure 7. As expected, the P1-Arg fits into the S1 binding pockets of both trypsin and K60E, forming hydrogen bonds with Ser190, Gly219, and Asp226 and van der Waals contacts with the pocket walls. Furthermore, the catalytic residue Ser195 is adjacent to the scissile bond; the oxyanion hole interactions are present, and the NH group of the P2′ residue forms a hydrogen bond with the backbone carbonyl group of Phe41 in both complexes. However, striking differences between the complexes are found at the S1′ subsite. Whereas in trypsin the P1′-Arg forms only hydrogen bonds between one guanidinyl nitrogen and two backbone carbonyl groups, K60E forms a salt bridge between the P1′-Arg and Glu60 with two hydrogen bonds. The distance between the hydrogen bond donor and acceptor is less than 3.0 Å in this salt bridge. Moreover, the terminal P3′-NH_2 group of Bz-Arg-Arg-Ala-NH_2 also forms a hydrogen bond with Glu60. The interaction energy of the most stable complex of K60E is −141 kcal/mol, which is much greater than the interaction energy of −134 kcal/mol for the trypsin complex.

A second calculation addressed the question of why K60E accepts P1′-Arg significantly better than P1′-Lys. The complex between K60E and Bz-Arg-Lys-Ala-NH_2 exhibits strong similarities with the K60E−Bz-Arg-Arg-Ala-NH_2 complex (Figure 7). However, only one hydrogen bond forms between Glu60 and the ϵ-amino group of P1′-Lys. This observation may explain the higher specificity of K60E for P1′-Lys.

Summary. Our data clearly show that engineering of the S′ subsite of trypsin is a suitable tool for designing site-specific proteases. Acyl transfer studies and peptide cleavage experiments consistently demonstrate that trypsin K60E has a strong S1′ specificity for Arg. Consequently, K60E preferably cleaves between Arg−Arg bonds. Molecular modeling studies suggest that this specificity is due to a salt bridge between Glu60 and P1′-Arg. In addition, modulation of trypsin specificity is not limited to the S1 and S1′ subsites. Indeed, we have recently altered the specificity of the S2′ site (T. Kurth, D. Ullmann, H.-D. Jakubke, and L. Hedstrom, unpublished results). Thus, we can now manipulate the specificity of three trypsin subsites, and it may be possible to design a protease which recognizes a specific three-residue sequence.

ACKNOWLEDGMENT

We thank Doris Haines for technical assistance.

REFERENCES

1. Ballinger, M. D., Tom, J., and Wells, J. A. (1995) *Biochemistry 34*, 13312−13319.
2. Ballinger, M. D., Tom, J., and Wells, J. A. (1996) *Biochemistry 35*, 13579−13585.
3. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
4. Carter, P., and Wells, J. A. (1991) *Science 237*, 394−399.
5. Corey, D. R., Willett, W. S., Coombs, G. S., and Craik, C. S. (1995) *Biochemistry 34*, 11522−11527.
6. Willett, W. S., Brinen, L. S., Fletterick, R. J., and Craik, C. S. (1996) *Biochemistry 35*, 5992−5998.
7. Walsh, K. A., and Wilcox, B. E. (1970) *Methods Enzymol. 19*, 31−69.
8. Fersht, A. R., Blow, D. M., and Fastrez, J. (1973) *Biochemistry 12*, 2035−2041.
9. Schellenberger, V., Turck, C. W., Hedstrom, L., and Rutter, W. J. (1993) *Biochemistry 32*, 4349−4353.
10. Schellenberger, V., Turck, C. W., and Rutter, W. J. (1994) *Biochemistry 33*, 4251−4257.
11. Hedstrom, L., Szilagyi, L., and Rutter, W. J. (1992) *Science 255*, 1249−1253.
12. Hedstrom, L., Perona, J., and Rutter, W. J. (1994) *Biochemistry 33*, 8757−8763.
13. Kurth, T., Ullmann, D., Jakubke, H.-D., and Hedstrom, L. (1997) *Biochemistry 36*, 10098−10104.
14. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.

15. Ullmann, D., and Jakubke, H. D. (1994) *Eur. J. Biochem. 223*, 865−872.
16. Pennington, M. W., and Thornberry, N. A. (1994) *Pept. Res. 7*, 72−76.
17. Perona, J. J., Hsu, C. A., Craik, C. S., and Fletterick, F. J. (1993) *J. Mol. Biol. 230*, 919−933.
18. Morris, G. M., Goodsell, D. S., Huey, R., and Olson, A. J. (1996) *J. Comput.-Aided Mol. Des. 10*, 293−304.
19. Goodsell, D. S., Morris, G. M ., and Olson, A. J. (1996) *J. Mol. Recognit. 9*, 1−5.
20. Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−290.
21. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Eng. 8*, 127−134.
22. Latt, S. A., Auld, D. S., and Vallee, B. L. (1972) *Anal. Biochem. 50*, 57−62.
23. Matayoshi, E. D., Wang, G. T., Krafft, G. A., and Erickson, J. (1989) *Science 247*, 954−958.
24. Craik, C. S., Largeman, C., Fletcher, T., Rocziak, S., Barr, P. J., Fletterick, R. J., and Rutter, W. (1985) *Science 228*, 291−297.
25. Schellenberger, V., and Jakubke, H.-D. (1991) *Angew. Chem., Int. Ed. Engl. 30*, 1437−1449.

BI980842Z